Developmental Cell, Vol. 6, 597–606, April, 2004, Copyright ©2004 by Cell Press

Altering the Expression in Mice of Genes by Modifying Their 3′ Regions

Technique

Masao Kakoki,[1] Yau-Sheng Tsai,[1] Hyung-Suk Kim,[1] Seigo Hatada,[1] Dominic J. Ciavatta,[1] Nobuyuki Takahashi,[1] Larry W. Arnold,[2] Nobuyo Maeda,[1] and Oliver Smithies[1,*]

[1]Department of Pathology and Laboratory Medicine
[2]Department of Microbiology and Immunology
University of North Carolina
Chapel Hill, North Carolina 27599

Summary

Polymorphic differences altering expression of genes without changing their products probably underlie human quantitative traits affecting risks of serious diseases, but methods for investigating such quantitative differences in animals are limited. Accordingly, we have developed a procedure for changing the expression in mice of chosen genes over a 100-fold range while retaining their chromosomal location and transcriptional controls. To develop the procedure, we first dissected the effects in embryonic stem (ES) cells of elements within and downstream of the 3′ untranslated region (UTR) of a single copy transgene at the *Hprt* locus. As expected, protein expression varied with the steady-state level and half-life of the mRNA. The rank order of expression with various tested 3′ regions is the same in ES cells, and in cardiomyocytes and trophoblastocytes derived from them. In mice having two functionally different native genes with modified 3′UTRs, the desired expression was obtained.

Introduction

The vast extent of human genetic variability is becoming increasingly obvious as the human genome project progresses (Lander et al., 2001; Venter et al., 2001). It is therefore reasonable to expect that common polymorphic alleles will exist that dictate expression of the products of many genes at different intrinsic levels without changing the products themselves or the location and timing of their expression. These considerations indicate the need to know the effects of genetic variations that change the level of expression of genes, and particularly of those relevant to complex multifactorial diseases.

One method of seeking this knowledge is to generate transgenic animals by the injection of DNA into the male pronucleus of one-cell murine embryos. However, DNA incorporated into the mouse germline in this way is integrated randomly and in unpredictable copy numbers. Moreover, the level of gene expression in the resulting mice frequently fails to correlate with the number of copies of the integrated genes and is often markedly influenced by the site of integration (Hatada et al., 1999). As a means of overcoming some of these problems, we previously developed a "gene titration" method in which homologous recombination in embryonic stem (ES) cells is used to delete or duplicate a chosen gene at its normal chromosomal location without changing the sequence of the gene product or of nearby *cis*-acting control elements (Smithies and Kim, 1994). Animals generated from the ES cells modified in this way have one, two (wild-type), three, or four functional copies of the gene of interest. A linear and strictly proportional relationship between the number of copies of a gene and the steady-state level of its product was observed in mice resulting from several experiments of this type (Krege et al., 1997; Oliver et al., 1998). This agrees with previous observations in humans that heterozygotes for a completely nonfunctional mutation usually have close to 50% normal levels of gene product (Harris, 1970), while trisomic individuals have close to 150% normal levels (Epstein, 1989). Fifty percent expression is consequently regularly observed at the mRNA level in mice heterozygous for a gene deletion "knock out." However, in the three- and four copy animals, the chromosome with the duplicated gene sometimes produces less than twice the product of the normal chromosome (Kim et al., 2002). This lack of proportionality is likely because in these cases the duplicated region does not include *cis*-acting regulatory elements that are important for normal transcription levels. Increasing the extent of the duplicated region can in principle overcome this problem, but this solution is not always applicable because of the difficulty of obtaining long duplications and because of complications introduced when the duplication includes an adjacent gene. Clearly, there is need for a method of changing the expression of a gene at its own chromosomal location, preferably over a wider range than can be achieved by deletion and duplication, while still retaining the controlling elements (promoters, enhancers, introns, etc.) and the effects of the chromosomal location that influence transcription of the wild-type gene.

Accordingly, we have developed a genetic procedure for varying the level of expression of a target gene which comes close to fulfilling this need. The procedure is based on the results of many previous studies showing that the 3′-untranslated region (3′UTR) of a gene influences the stability of its mRNA (Ross, 1995). To develop the new procedure, we first constructed a quantitative reporter system using a green fluorescent protein transgene inserted by homologous recombination into the hypoxanthine-guanine phosphoribosyltransferase (*Hprt*) locus as a single copy. We then studied the effects on the expression of the reporter of using natural or modified 3′UTRs with or without their downstream GU/U-rich sequence elements (GREs) (Gil and Proudfoot, 1987). We show that the steady-state level of expression of the reporter gene can be varied over more than a 100-fold range by modifying its 3′ region. The relative expression achieved in ES cells with different 3′ regions is preserved in two types of differentiated cells derived from the ES cells. Data are presented confirming that the desired levels of expression, one an increase and the other a decrease, were obtained in two lines of mice

*Correspondence: jenny_langenbach@med.unc.edu

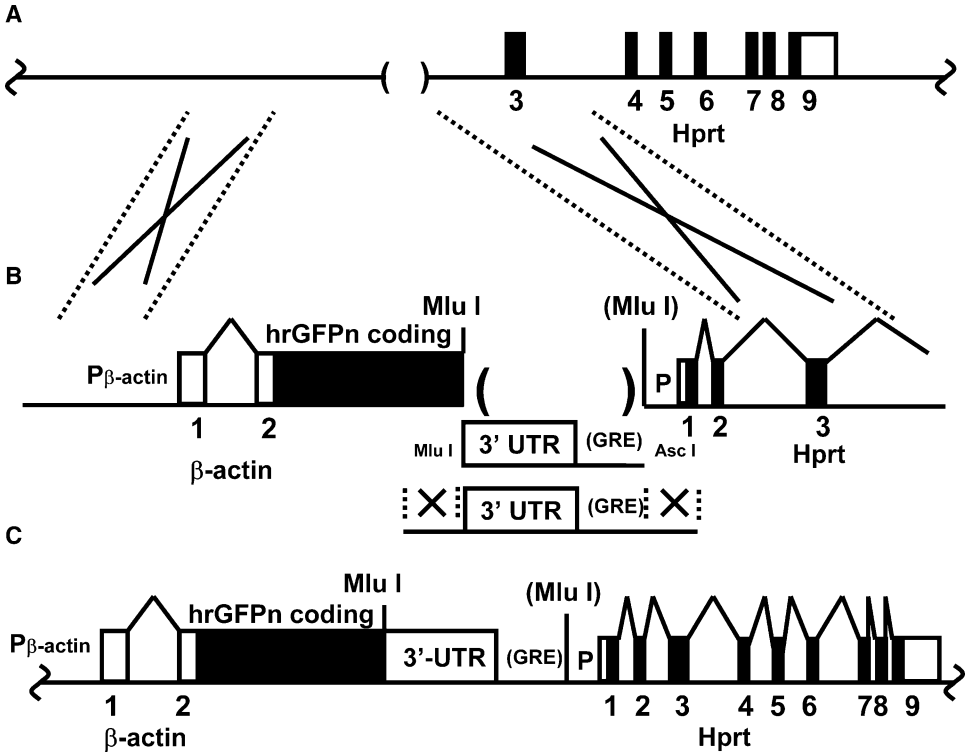

Figure 1. A Single Copy Transgene Strategy for Testing the Effects of 3′ Regions on the Expression of a Humanized Renilla Green Fluorescence Protein with a Nuclear Localization Signal (*hrGFPn*)

(A) The target locus, into which the transgene is introduced by homologous recombination (Doetschman et al., 1987), is the *Hprt* gene of BK4 ES cells; it lacks the promoter and exons 1 and 2. (B) The targeting vector contains a human β-actin promoter (Pβ-actin) and two noncoding β-actin exons (1,2), the *hrGFPn* coding sequence with an MluI site just 3′ to the stop codon. Different 3′UTRs with or without their GU/U-rich sequence elements (GRE) are introduced into this MluI site by ligation using MluI and AscI sites on the 5′ and 3′ ends of the test regions, or by recombination in *E. coli* (Zhang et al., 1998). (C) The resulting locus after homologous recombination and selection in hypoxanthine-aminopterin-thymidine (HAT) media. The *Hprt* gene has a promoter (P) and all 9 exons and the locus now contains a single copy transgene. Mlu and (Mlu) indicate MluI sites that are intact or have been destroyed by ligation.

derived from ES cells in which the 3′UTRs of two functionally different native genes had been modified. Thus modification of the 3′ region of a gene is a potent tool for changing its expression in mice in a predictable fashion while retaining its chromosomal position, promoters, and introns.

Results

The Test System

Figure 1 illustrates diagrammatically the homologous recombination and HAT selection (Doetschman et al., 1987) used to introduce into the *Hprt* locus of *Hprt*-negative ES cells a single copy of a reporter transgene, *hrGFPn* (humanized renilla Green Fluorescence Protein with a nuclear localization sequence), having a 3′ region to be tested. Use of this procedure allowed us to compare many different 3′ regions for their effects on gene expression. Because each construct is tested as a single copy transgene in a fixed locus (*Hprt*), the observed fluorescences enable rigorous comparisons of the effects on expression of the tested 3′ regions.

The 3′ Regions Tested and the Rank Order of Expression

Figure 2 illustrates, in the rank order of the levels of expression in ES cells that they confer, 28 natural or modified 3′ regions that we have tested, together with histograms from FACS analyses of the corresponding ES cells and the parental ES cell, BK4. Note that, except as otherwise indicated, all the elements included in each of the tested 3′ regions are from the gene named in its title. For example, "β-globin +" in line 6 of the figure indicates that the tested 3′ region included the β-globin 3′UTR with its own natural poly(A) addition signals plus the GRE from the same gene; "β-globin" on line 13 indicates that the 3′ region included the β-globin 3′UTR with its own poly(A) addition signals but did not include the GRE. (The 5′ and 3′ primers used to generate the various 3′ regions are included in Supplemental Table S1 [http://www.developmentalcell.com/cgi/content/full/6/4/597/DC1].) The data are presented in rank order so that the relative "strengths" of the different 3′ regions can be seen at a glance, and our various figures can be readily compared.

The histograms presenting the results of the FACS analyses of cells (at least 10,000 cells per analysis) were obtained by trypsinizing culture dishes with many colonies. As can be seen, the counts of cells (linear vertical axis) having different intensities of fluorescence (logarithmic horizontal axis) are monophasically and essentially normally distributed. The mean fluorescences of the modified ES cells (the numbers in the histogram

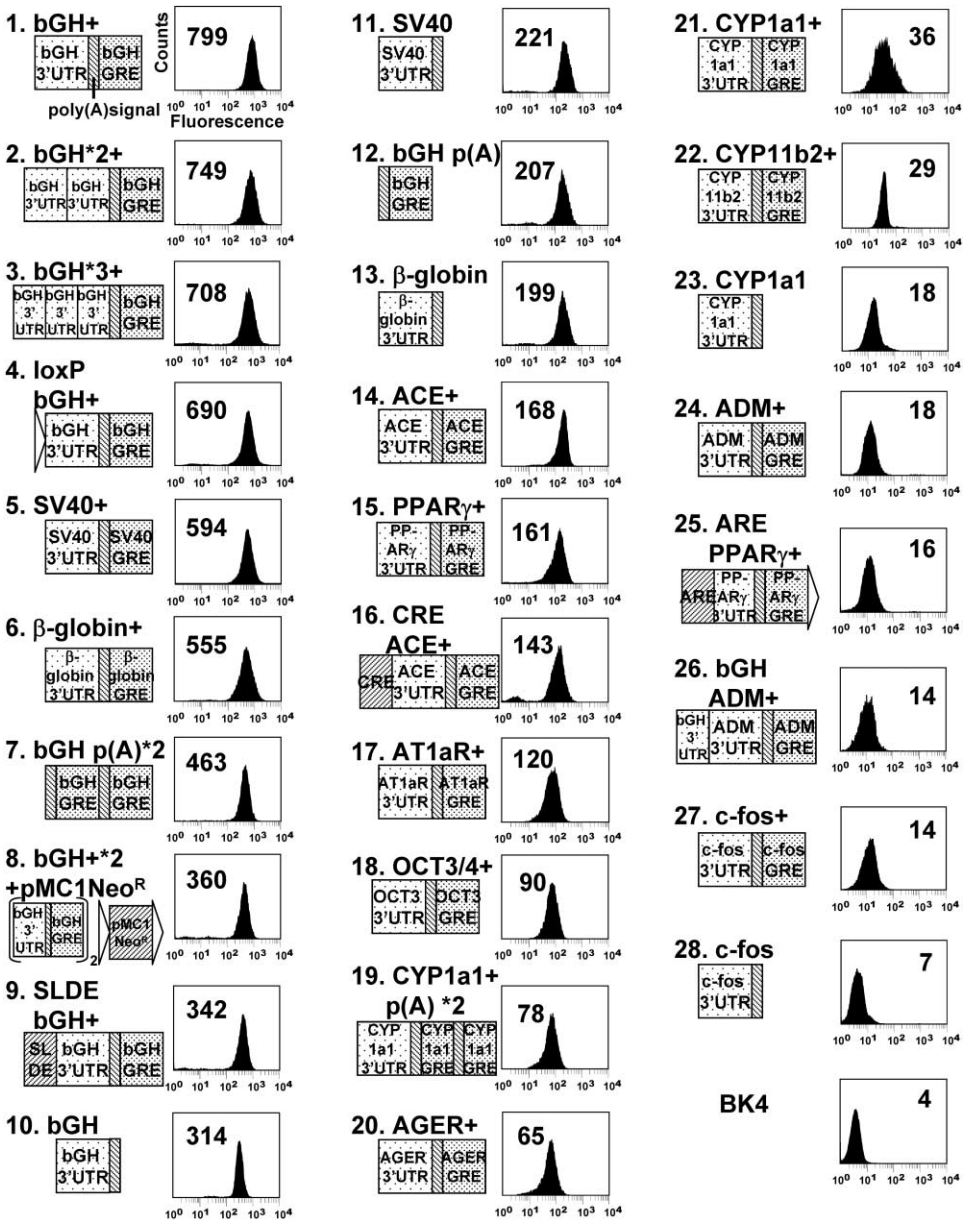

Figure 2. Diagrammatic Representation of the 28 3′ Regions Tested in This Study and the Results of FACS Analysis of the Modified ES Cells

In the diagrams, 3′UTR indicates sequences downstream of the stop codon up to but not including the poly(A) signal; GRE, a GU/U-rich element; bGH, bovine growth hormone; +, a 3′UTR with its own GRE; *2 and *3, duplicated or triplicated sequences; SLDE, stem-loop destabilizing element; SV40, the late gene of SV40; pMC1Neo^R, MC1 promoter-driven neomycin phosphotransferase gene; AT1aR, angiotensin 1a receptor; ACE, angiotensin converting enzyme; CRE, minimal C-rich element of human α-globin; ADM, adrenomedullin; OCT-3/4, octamer binding protein-3/4; AGER, advanced glycation end product receptor; PPARγ, peroxisome proliferator-activated receptor γ; ARE, AU/U-rich element of mouse *c-fos* gene.

3′UTR sequences are lightly stippled. Poly(A) addition signals are slashed left slanting. GREs are heavily stippled. Various unique sequences are slashed right slanting. Open triangles are loxP sequences. The constructs are numbered for reference in other figures and in the text. The sequences in a given 3′ region are from the gene named in its title, except when indicated otherwise.

The histograms show the distribution of fluorescence intensity in trypsinized ES cells (n > 10,000). The vertical scales of the FACS histograms are linear and represent counts of the numbers of cells at each level of fluorescence. The horizontal scales are logarithmic and cover fluorescence intensity from 1 to 10^4. The numbers on the histograms are the mean fluorescence intensity.

panels) range from (for bGH+, line 1) approximately 200X the background fluorescence of untransformed BK4 ES cells to 2X (for c-*fos*, line 28). We describe below how the data in Figure 2 can be used to select a suitable 3′ region for altering the expression of a chosen target gene in a desired manner.

Systematic Effects of 3′ Region Modifications

The 3′ region constructs used to generate the data in Figure 2 were designed primarily to enable the expression of a target gene to be varied in a graded manner between the limits of the maximum obtainable (with bGH+) to the minimum (with c-*fos*). However, the data

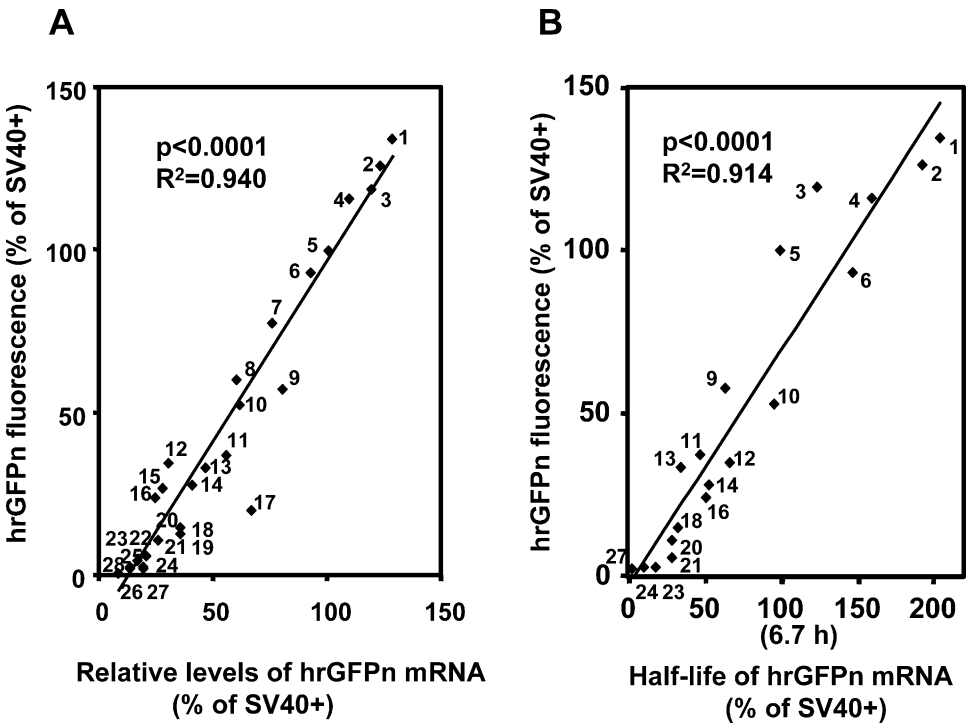

Figure 3. The hrGFPn Fluorescence with Each Construct Correlates with Its mRNA Level and the Half-Life of Its mRNA

(A) Correlation between the steady-state level of *hrGFPn* mRNA and the mean intensity of hrGFPn fluorescence with constructs having different 3′ regions. hrGFPn mRNA levels are normalized to β-actin mRNA and fluorescence is relative to SV40+ as 100%. The numbering follows that used in Figure 2.

(B) Correlation between the half-lives of hrGFPn mRNA and fluorescence intensity. Both are relative to that of SV40+ as 100%. Half-lives were determined from the values of *hrGFPn* mRNA/β-*actin* mRNA 0, 1, 3, 6, 12, and 24 hr after adding 10 μg/ml actinomycin D to the culture medium.

also allow several systematic effects to be noted. Thus we confirm previous work (Gil and Proudfoot, 1987) showing that inclusion or omission of the GRE downstream of the poly(A) addition signals of a 3′UTR allows expression to be altered by a factor of between two and three (compare lines 1 and 10, 5 and 11, 6 and 13, 21 and 23, 27 and 28). We also confirm previous observations (Shaw and Kamen, 1986) that expression is markedly reduced when an AU-rich element (ARE), such as occurs in the natural *c-fos* 3′UTR, is placed between the stop codon and a 3′UTR (lines 15 and 25 in Figure 2 and Figure 3, below). Likewise, we confirm observations (Brown et al., 1996) that a Stem Loop Destabilizing Element (SLDE) reduces expression (lines 1 and 9). In contrast, the loxP stem loop has almost no effect (lines 1 and 4), an observation with important implications for future work (see discussion). We find that duplicating or triplicating a 3′UTR without its poly(A) addition signals does not affect expression (lines 1, 2, and 3). However, expression is doubled by duplicating a 3′ poly(A) addition signal and GRE (lines 7 and 12, 19 and 21).

Quantitative Analysis of mRNA Expression and Stability

To determine whether the intensity of fluorescence obtained with a given construct (a measure of expression at the protein level) parallels the level and half-life of the corresponding mRNA, we used quantitative RT-PCR to measure steady-state mRNA levels and followed the decrease in level when transcription was inhibited with actinomycin D. Figure 3A shows that the fluorescence obtained with all 28 constructs having different 3′ regions is linearly correlated ($p < 0.0001$; $R^2 = 0.940$) with the level of *hrGFPn* mRNA. There are no significant differences in the rank order of the 3′ regions when measured in the two ways. Thus the expression of *hrGFPn* at the protein level, as judged by fluorescence, is quantitatively proportional to expression of the *hrGFPn* gene at the mRNA level. Figure 3B demonstrates that the steady-state level of *hrGFPn* mRNA with different 3′ regions is, as expected, positively correlated ($p < 0.0001$; $R^2 = 0.914$) with the relative half-lives of their mRNAs normalized to the half-life of β-*actin* mRNA.

Effects of 3′UTR Modifications in Differentiated Cells

The foregoing data demonstrate that systematic alterations of the expression in ES cells of a single copy transgene can be achieved by changing its 3′ region. For this form of gene modification to be generally useful for systematically altering the expression of a target gene in a mouse, the effects seen in ES cells must also be seen in differentiated cells derived from these cells.

To determine whether this requirement is met, we removed the ES cells from the feeder layer cells that maintain them in the undifferentiated state, and cultured them under conditions leading to a variety of differentiated cells. We chose cardiomyocytes as one type of differentiated cell in which to ask whether expression in a differentiated cell is comparable to that in the ES cell from which it was derived. Cardiomyocytes are useful for this purpose since they are regularly produced when ES cells differentiate and are easily recognized by their spontaneous beating. A second type of differentiated cell that regularly arises in our cultures is a trophoblast-like cell, easily recognized by its large size and one or more prominent nuclei. The cardiomyocytes and trophoblastocytes were respectively immunoreactive to antibodies for troponin T and placental lactogen-1 which are specific markers for these two cell types (bottom two panels of Figure 4). Figure 4 compares the rank order of the fluorescence of the modified ES cells for seven of our tested series of 3′ regions with that from cardiomyocytes and trophoblastocytes derived from these ES cells. The rank order of the seven 3′ regions in the parent ES cells and in the two types of differentiated cells generated from them is identical.

Decreasing Gene Expression

Changing the 3′ region of a target gene by replacement with a different 3′ region is likely, in most cases, to yield a useful increase or decrease in expression. However, in some cases, replacing the natural 3′ region of a gene with a different 3′ region may remove important regulatory sequences. An example would be iron-responsive elements in the 3′UTR of the transferrin gene (Binder et al., 1994). Accordingly, our experiments also included tests of ways of modifying the expression of a gene by the *addition* of suitable sequences while retaining most or all of the gene's natural 3′ region. Thus we tested the effects of adding an element (an ARE from *c-fos*) (Shaw and Kamen, 1986) to the 3′UTR of the *PPAR*γ gene in order to decrease its expression (compare lines 15 and 25 in Figure 2), as discussed in more detail below. Likewise, the addition of an SLDE from the granulocyte colony stimulating factor gene (Brown et al., 1996) decreased expression of *bGH* to about half (lines 1 and 9). However, the loxP sequence, which has a stem-loop structure and YAT sequences in its loop similar to those in SLDE, decreased fluorescence by only 15% (lines 1 and 4). Deletion of the GRE downstream of an existing 3′UTR decreased expression to about half or one-third in the case of several 3′UTRs (lines 1 and 10, 5 and 11, 6 and 13, 21 and 23, 27 and 28). Removal of the 3′UTR sequences upstream of a poly(A) signal can also decrease expression (to one-quarter in the case of *bGH*; lines 1 and 12), but this is at the cost of losing information in the deleted region.

Increasing Gene Expression

We had more difficulty in finding ways of increasing the expression of genes while retaining most of their natural 3′UTRs. The addition of minimal C-rich stabilizing elements (CRE) from the β-*globin* gene (Holcik and Liebhaber, 1997) into the 3′UTR after the stop codon of the *Ace* gene did not increase the fluorescence (lines 14 and 16). Similarly, introducing a portion of the *bGH* 3′UTR from between the stop codon and poly(A) signal into the adrenomedullin (*Adm*) 3′ region did not increase fluorescence (lines 24 and 26), even though omitting the same sequence from the bGH 3′UTR diminished fluorescence considerably (lines 1 and 12). Nevertheless, we found that a potentially general strategy for increasing expression without introducing sequences not already present in the 3′ region is to duplicate the poly(A) signal and accompanying GRE. Thus, a simple poly(A) signal and GRE immediately after the stop codon of our reporter gene gives good expression (line 12 in Figure 2), and duplicating this sequence increased expression about 2-fold (line 7). Likewise, duplicating the poly(A) signal and GRE of the *CYP1a1* gene gave a 2-fold increase in expression (lines 19 and 21).

Designing the 3′ Region for a Target Gene

In order to design a modification to obtain a desired change in the expression of a chosen target gene, it is necessary to compare the "strength" of the natural 3′UTR and GRE from the target gene with that of previously tested 3′ regions. We illustrate this aspect of our system with the gene coding for the angiotensin type 1a receptor (*AT1aR*), whose expression we wished to increase in the mouse, and with the gene coding for the peroxisome proliferator-activated receptor γ (*PPAR*γ), whose expression we wished to decrease. Both of these target genes have 3′UTRs and GREs that lead to relatively low levels of expression (20% of SV40+ for *AT1aR*, lines 5 and 17; and 27% for *PPAR*γ, lines 5 and 15). To increase the expression of *AT1aR*, we therefore chose to replace its natural 3′ region with a construct (Figure 5A, and line 8 in Figure 2) that included two copies of the 3′UTR and GRE from the *bGH* gene plus a selectable *Neo*^{R} marker flanked by loxP sequences. To decrease the expression of the *PPAR*γ gene, we decided to insert into its natural 3′UTR an ARE from *c-fos* plus the loxP sequence that remains after the action of *Cre*-recombinase (Figure 5C, and line 25 in Figure 2).

The results of incorporating these planned targeted modifications into our test system led us to expect that *AT1aR* expression should be increased to about 3 X wild-type in homozygotes (compare lines 8 and 17 in Figure 2) and 2 X in heterozygotes, while *PPAR*γ expression should be reduced to about 0.1 X wild-type in homozygotes (compare lines 15 and 25) and 0.55 X in heterozygotes.

Tests in Animals

We have generated animals from the ES cells that were modified in the ways illustrated in Figure 5A and 5C, and have used them to test the overall effectiveness of our procedure. Figure 5B shows that animals heterozygous for the *bGH* 3′UTR targeted modification of the *AT1aR* gene have in vivo expression of the *AT1aR* gene in kidney and heart (two tissues in which the expression can be measured with precision) increased to about 1.8 X of wild-type. Animals homozygous for the *c-fos* 3′UTR targeted modification of the *PPAR*γ have in vivo expression of the *PPAR*γ gene in adipose tissue (the primary site of its function) reduced to about 0.1 X of wild-type. Expression is also decreased to similar extents in heart,

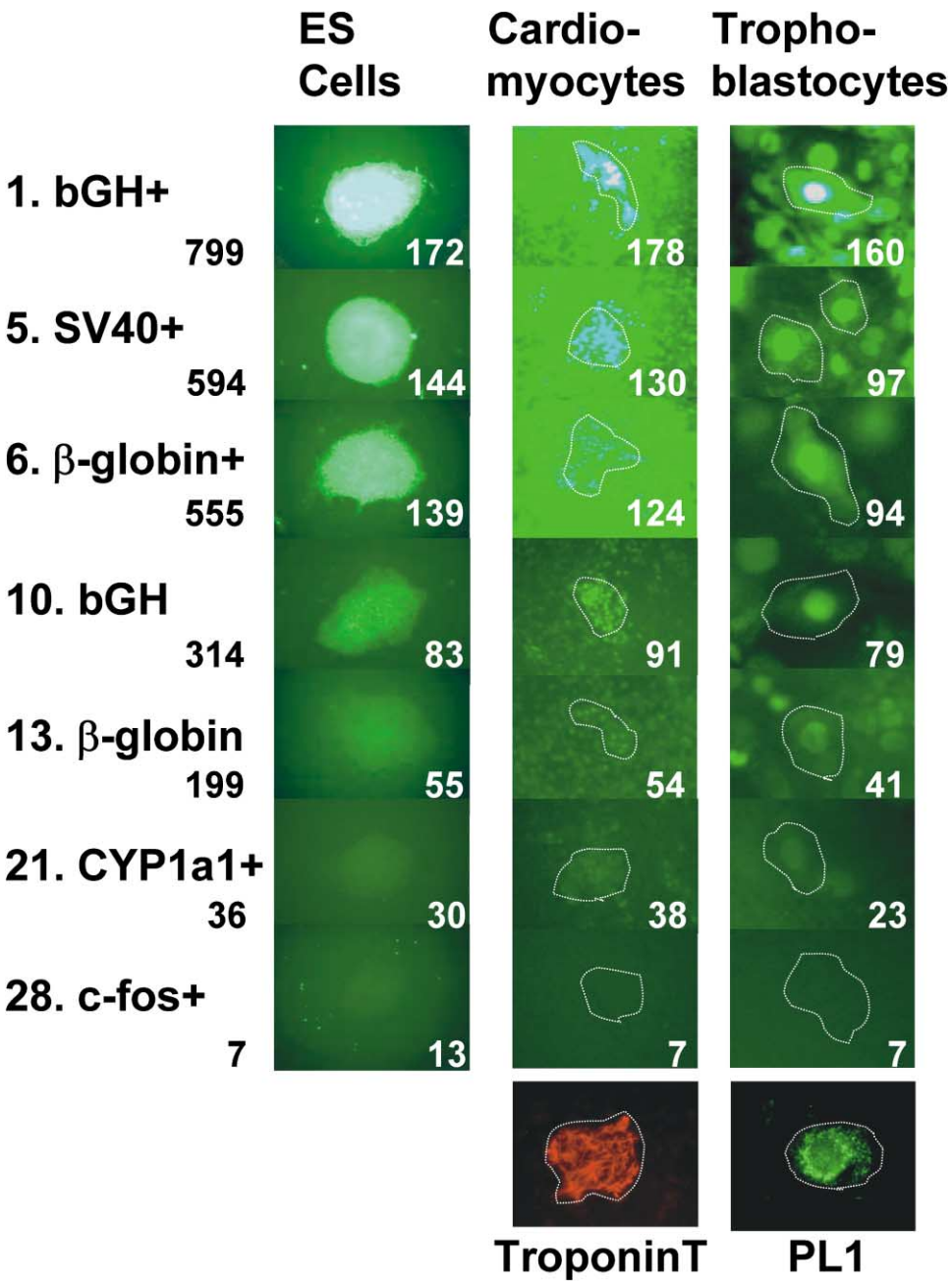

Figure 4. Digital Images and Intensities of the Green Fluorescence from Seven Differently Modified ES Cells and from Cardiomyocytes and Trophoblastocytes Derived In Vitro from Them 10 Days after Plating on Laminin-Coated Dishes

The dotted lines indicate the borders of the cardiomyocytes and trophoblastocytes. The labels for the cells follow those used in Figure 2, and the black numbers immediately next to the labels record the average fluorescence of FACS sorted cells taken from Figure 2. The white numbers in each image panel indicate the fluorescence intensities of the nuclei of the imaged cells corrected for background fluorescence. Cardiomyocytes were identified by their beating; they express troponin T, a characteristic marker of this cell type, as shown in the left bottom panel. Trophoblastocytes were identified by their characteristic morphology; they express placental lactogen 1 (PL1), a characteristic marker of this cell type, as shown in the right bottom panel.

liver, and kidney. Thus the results in animals agree very well with design expectations since the expressions of two functionally quite different target genes were modified in opposite directions, as planned, in the several tested tissues, and to degrees close to those predicted by the prior experiments in ES cells.

Discussion

Much effort has been expended in the past on manipulations to control the level of gene expression through factors that affect the two synthetic processes, transcription and translation. Yet the steady-state levels of

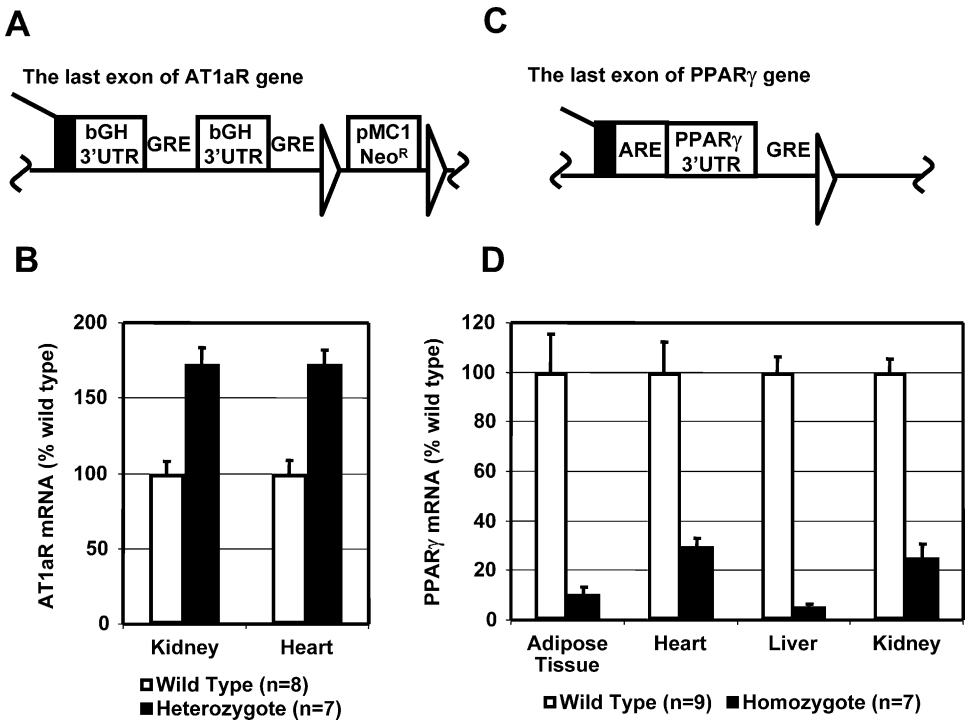

Figure 5. Generation of Animals with Genes Modified in Their 3′ Regions

(A) 3′ region used to increase expression of the angiotensin 1a receptor (*AT1aR*) in an animal.

(B) Kidney and heart expression of the modified *AT1aR* gene (black bars) relative to wild-type as 100% (white bars).

(C) 3′ region used to decrease expression of the peroxisome proliferator-activated receptor γ (*PPARγ*).

(D) Adipose tissue, heart, liver, and kidney expression of the modified *PPARγ* gene (black bars) relative to wild-type (white bars).

The 3′ regions illustrated in (A) and (C) are as in lines 8 and 25 of Figure 2.

all cell components are proportionately as much affected by changes in the degradation of the mRNA and protein products as they are by changes in their synthesis. For this reason, we chose to develop a system for modifying gene expression that takes advantage of changes affecting mRNA stability (Hargrove and Schmidt, 1989). Among the factors affecting mRNA degradation, the 3′UTR and its poly(A) sequence play major roles. Thus the first step in the decay of many mRNAs is deadenylation followed by decapping (Wang and Kiledjian, 2001). Since poly(A) binding proteins protect mRNAs from deadenylation, they also play an important role in mRNA stabilization (Sachs and Davis, 1989). Additionally, it has been demonstrated that GREs downstream of the poly(A) signal(s) are required for efficient polyadenylation of many mammalian transcripts (Gil and Proudfoot, 1987). Indeed, cleavage and polyadenylation of transcripts are coordinate functions controlled by polyadenylation specificity factors that bind to the poly(A) signal, and by cleavage stimulation factors that bind to the GRE (de Vries et al., 2000).

Other elements affect mRNA stability. For example, Shaw and Kamen demonstrated that the introduction of a 51-nucleotide AU-rich element (ARE) from the short-lived mRNA of a human lymphokine gene coding for granulocyte-monocyte colony stimulating factor reduced the steady-state cytoplasmic mRNA levels of a rabbit β-globin gene without affecting the rate of its transcription (Shaw and Kamen, 1986). The effect was reproducible in three different cell lines tested. Later studies showed that the presence in a 3′UTR of AREs that include the AUUUA pentanucleotide motif confers an unstable property on the mRNA of protooncogenes (Brewer and Ross, 1988), immune regulators (Chen et al., 1995), and β-adrenergic receptors (Huang et al., 1993). During ARE-mediated decay, shortening of the poly(A) tail precedes the degradation of the mRNA body. Our present experiments demonstrate that these various features of 3′UTRs and their downstream GREs can be used to alter the expression of a target gene systematically.

The first set of experiments examined the effects on expression in ES cells of changing the 3′UTR and GRE of a reporter transgene for a fluorescent protein (*hrGFPn*) driven by a β-actin promoter and inserted into the X-linked *Hprt* locus. Expression, judged by fluorescence or quantification of the mRNA level by RT-PCR, was controlled over more than a 100-fold range by changing the 3′ region (Figures 2 and 3). Single-cell fluorescence (a measure of the protein concentration) and mRNA levels were directly and linearly correlated ($R^2 = 0.940$, $p < 0.0001$, Figure 3A). Likewise, mRNA levels and their half-lives were highly correlated ($R^2 = 0.914$, $p < 0.0001$, Figure 3B).

Several systematic observations made in the course of the first set of experiments require comment. First,

we find that the final levels of expression decrease to about one-half to one-third when a 3′UTR is introduced into the genome without its downstream GRE. Second, we confirm previous reports that insertion of AREs (Shaw and Kamen, 1986) or SLDEs (Brown et al., 1996) into a 3′UTR decreases protein expression in parallel with a decrease in mRNA steady-state level. Third, we find that the insertion of loxP sequence into the *bGH* 3′ region scarcely changes expression. This observation is of particular interest since it suggests that tissue-specific or stage-dependent changes in expression from high to low or low to high should be relatively easy to obtain by using two 3′ regions in a single targeting construct (for example, loxP-bGH-loxP-c-*fos* or loxP-c-*fos*-loxP-bGH). Fourth, we find that although removal of most of the *bGH* 3′UTR between the stop codon and poly(A) signal markedly reduced expression (lines 1 and 12 in Figure 2), addition of the same sequence to the 3′UTR of the *Adm* gene did not change expression (lines 24 and 26). Similarly, although Holcik and Liebhaber (1997) report that removal of a C-rich element from the human α-*globin* gene reduced its expression, we found that addition of the same sequence to the 3′UTR of the *Ace* gene did not change expression (lines 14 and 16). These observations suggest that the effects of adding "foreign" stabilizing elements from a different gene may be context dependent, whereas the addition of destabilizing elements appears less sensitive to context. However, we found, fifth, that duplicating the poly(A) addition signal and GRE of the 3′ region of the *CYP1a1* gene doubled expression (lines 19 and 21). Likewise, duplicating a simple poly(A) addition signal and GRE (line 12) also doubled expression (line 7).

Our second set of experiments asked whether the relative effects on expression of the different 3′ regions was the same in several types of cell. The relevance of this to using the procedure for whole animal experiments is obvious; the procedure is likely to be more simple to apply if the relative "strengths" of the 3′ regions are the same in all cell types. Figure 4 presented data showing that the rank order of expression of seven tested 3′ regions was the same in three quite different cell types (ES cells, cardiomyocytes, and trophoblastocytes). This parallels the observations by Shaw and Kamen on the reproducible effects of AREs in three different cell types (Shaw and Kamen, 1986).

An important caveat is required when considering replacing the 3′ region of a target gene that has a complex 3′UTR with the 3′ region from another gene, namely the possibility that specialized regulatory functions intrinsic to the wild-type gene will be lost. Our experiments suggest that this problem can be avoided with modifications aimed at decreasing expression by inserting destabilizing elements without removing specialized sequences from the natural 3′UTRs of the target gene. Likewise, duplicating the poly(A) signal and the GRE of a chosen gene appears to be a promising way of progressively increasing expression without either removing any sequences or introducing foreign sequences not already present in the 3′ region. However, it is important to recognize that currently none of our tested 3′ regions yields more expression than the *bGH* 3′UTR with its poly(A) addition signal and GRE. In the opposite direction, none has yielded expression less than the *c-fos* 3′UTR with its poly(A) addition signal but without a GRE. Attempts to use our procedure to increase the expression of a gene with a 3′ region comparable in strength to that from *bGH*, or to decrease the expression of a gene with a 3′ region as weak as that from *c-fos* should therefore be approached with caution.

A third set of experiments tested the ability of our overall strategy to increase to desired levels the expression in animals of a chosen gene (*AT1aR*) and decrease the expression of another (*PPARγ*). In vitro comparisons of the natural 3′ regions of these target genes with previously tested 3′ regions enabled the design of modifications to alter expression as desired. Animals derived from ES cells in which the target genes were so modified expressed the altered genes in several tested tissues at close to predicted levels. Particularly gratifying is the survival of animals homozygous for the planned low expression of *PPARγ*, since previous efforts to study the effects of low expression of this gene have been hampered by the failure of homozygous knockout animals to survive (Barak et al., 1999; Kubota et al., 1999).

In summary, we have demonstrated that modifying the 3′ region of a target gene provides a generally applicable way of increasing or decreasing its expression in a predictable fashion, while retaining its chromosomal location, promoters, and introns. We have compiled a panel of tested 3′ regions that can be used to vary expression over a range of more than 100-fold. We also describe ways of changing the expression of a gene of interest by modifying its 3′ region while still retaining features of the region that are special to the target gene. We illustrate the application of our procedure to predictably alter the expression of two target genes in the mouse genome.

Experimental Procedures

Targeting Vectors and Constructs

The *hrGFPn* 3′UTR test vector (Figure 1B) was made by inserting into our previously described *Hprt*- targeting vector (Bronson et al., 1996) a cDNA coding for a humanized Renilla green fluorescence protein with a nuclear localizing sequence (*hrGFPn*; Stratagene) driven by a human β-actin promoter sequence together with its first intron. An MluI site was introduced immediately 3′ to the stop codon of *hrGFPn*. Various 3′UTRs with or without the addition of their downstream GREs were cloned into this MluI site by ligation or by homologous recombination in *E. coli* (Zhang et al., 1998) for tests of their effects on *hrGFPn* expression. To facilitate determining the orientation of the UTRs in the resulting constructs, we introduced an MluI site on the 5′ end of the test UTRs and an AscI site on the 3′ end. All targeting constructs were linearized with PmeI before electroporation. Figure 2 illustrates the 3′ regions tested in the present study. Supplemental Table S1 lists the primers used for generating these 3′ sequences by PCR with mouse strain 129 genomic DNA as a template or cloned fragments from it. In some cases, pBS β-geo bpA (Goodwin and Rottman, 1992) was the template.

Cell Lines and Culture

The mouse embryonic stem (ES) cell line used in our study is a subclone (BK4) of E14TG2a, a cell line derived from 129/Ola mice with a deletion in its *Hprt* gene (Hooper et al., 1987) (Figure 1A). The ES cells were grown on murine embryonic fibroblasts in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% heat-inactivated fetal bovine serum, 10 μM 2-mercaptoethanol, and 2 mM L-glutamine. For the gene targeting experiments, $\sim 3 \times 10^7$ ES cells were electroporated in the presence of 15 μg of the targeting DNA in a cuvette having a gap size of 4 mm and an area of 160 mm^2 with a 1 s discharge from a 250 μF capacitor charged to 300V.

Homologous recombinants were selected in HAT medium (16 μg/ml hypoxanthine, 0.175 μg/ml aminopterin, and 4.8 μg/ml thymidine) for 12 days, at which time individual drug-resistant colonies were picked for expansion and verification by Southern analysis. For in vitro differentiation experiments, $\sim 3 \times 10^7$ ES cells were trypsinized and transferred to a 100 mm plastic bacterial culture dish containing supplemented DMEM. Two days later, the resulting embryonic bodies were transferred to laminin-coated six-well culture plates and cultured for 10 days, again in supplemented DMEM.

Immunocytochemistry

Cells grown on laminin-coated glass coverslips were fixed with 2% paraformaldehyde and permealized with 0.05% Triton X-100. After 30 min blocking with normal rabbit serum (1:100), diluted primary antibodies for troponinT and placental lactogen 1 were applied for 1 hr at 37°C. Following a 10 min wash with phosphate-buffered saline (PBS), goat anti-rabbit IgG antibody conjugated with FITC or TRITC (1:100, KPL) was applied for 1 hr at 37°C. The coverslips were rinsed in PBS and mounted in VectaShield mounting medium (Vector Laboratories).

Fluorescence Microscopy and Image Analysis

The fluorescence of cells was evaluated with an Olympus IMT-2 microscope equipped with a filter set for FITC (Exciter 480/40, Dichroic 505, Emitter 535/50 nm) or TRITC (Exciter 535/50, Dichroic 565, Emitter 610/75). Digitized images were produced with an Optronics TAC470 camera and Scion Image 1.62a software (N.I.H.). The exposure time was a quarter second for all pictures, so that fluorescence intensity could be quantitated by averaging the signal from pixels that cover cell nuclei and subtracting the average background signal.

Flow Cytometry Analysis

Undifferentiated ES cells were trypsinized and resuspended in 1 ml of supplemented DMEM, and filtered through a 30 μm polyester mesh (Miltenyi Biotech) to remove cell aggregates. Flow cytometry for fluorescence was with a Cytomation MoFlo Cytometer equipped with an argon laser for excitation at 488 nm. A minimum of 10,000 events were collected for each analysis. The GFP fluorescence was detected with a 530/30 nm bandpass filter. Data acquisition and analysis were with Cytomation Summit 3.1 software, which provides histograms and a numerical value for the average fluorescence of the sorted cells.

Quantitative Reverse Transcription-PCR

Expression of the mRNA for *hrGFPn, AT1aR*, and *PPARγ* was measured by precision quantitative reverse transcription-PCR with the Perkin-Elmer ABI 7700 Sequence Detection System as previously described (Kim et al., 2002). Primers for *hrGFPn* amplification were 5′-ATCCTGGTGTACCGCCTGAA-3′ (Fwd) and 5′-TGCTGGATGAA GTGGTACTC-3′ (Rev). The probe for *hrGFPn* detection was 5′-FAM-CAGCTGCCACATGCGCACCCT-Tamra-3′. Primers for *PPARγ* amplification were 5′-AGACATGAGCCTTCACCCC-3′ (Fwd) and 5′-AGA AGGAACACGTTGTCAGC-3′ (Rev). The probe for *PPARγ* detection was 5′-FAM-CAAGTCCTTGTAGATCTCCTGGAGC-Tamra-3′. The primers and probes used for AT1aR and β-actin amplification and detection were described previously (Kim et al., 2002). All reactions included β-actin as an internal standard. 5×10^6 ES cells or 100 mg of tissue was lysed in 0.2 or 1 ml lysis buffer made by diluting 2X RNA lysis buffer (PE Biosystems) with Ca^{2+}- and Mg^{2+}-free PBS. The lysed cell solution was stored at −20°C before use for RNA isolations with the ABI Prism 6700 automated nucleic acid workstation (PE Biosystems). All amplifications were duplicated.

Generation of Animals with Modified 3′ Regions for *AT1aR* and *PPARγ*

Using our standard gene-targeting protocols (Koller et al., 1989), we generated two strains of mice having experimentally modified 3′ regions to determine whether the *hrGFPn* transgene 3′ region assays in ES cells can predict the expression level of the gene of interest in an animal derived from these cells.

Acknowledgments

This work was supported by grants from the National Institutes of Health to O.S. and a fellowship program from Sankyo Foundation of Life Science to M.K.

Received: October 17, 2003
Revised: February 12, 2004
Accepted: February 13, 2004
Published: April 12, 2004

References

Barak, Y., Nelson, M.C., Ong, E.S., Jones, Y.Z., Ruiz-Lozano, P., Chien, K.R., Roder, A., and Evens, R.M. (1999). PPARγ is required for placental, cardiac and adipose tissue development. Mol. Cell *4*, 585–595.

Binder, R., Horowitz, J.A., Basilion, J.P., Koeller, D.M., Klausner, R.D., and Harford, J.B. (1994). Evidence that the pathway of transferrin receptor mRNA degradation involves an endonucleolytic cleavage within the 3′ UTR and does not involve poly(A) tail shortening. EMBO J. *13*, 1969–1980.

Brewer, G., and Ross, J. (1988). Poly(A) shortening and degradation of the 3′ A+U-rich sequences of human c-myc mRNA in a cell-free system. Mol. Cell. Biol. *8*, 1697–1708.

Bronson, S.K., Plaehn, E.G., Kluckman, K.D., Hagaman, J.R., Maeda, N., and Smithies, O. (1996). Single-copy transgenic mice with chosen-site integration. Proc. Natl. Acad. Sci. USA *93*, 9067–9072.

Brown, C.Y., Lagnado, C.A., and Goodall, G.J. (1996). A cytokine mRNA-destabilizing element that is structurally and functionally distinct from A+U-rich elements. Proc. Natl. Acad. Sci. USA *93*, 13721–13725.

Chen, C.Y., Xu, N., and Shyu, A.B. (1995). mRNA decay mediated by two distinct AU-rich elements from c-fos and granulocyte-macrophage colony-stimulating factor transcripts: different deadenylation kinetics and uncoupling from translation. Mol. Cell. Biol. *15*, 5777–5788.

de Vries, H., Ruegsegger, U., Hubner, W., Friedlein, A., Langen, H., and Keller, W. (2000). Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. EMBO J. *19*, 5895–5904.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature *330*, 576–578.

Epstein, C.J. (1989). Down Syndrome (Trisomy 21). In The Metabolic Basis of Inherited Disease, C.R. Scriver, A.R. Beaudet, W.S. Sly, and D. Valle, eds. (New York: McGraw-Hill), pp. 291–326.

Gil, A., and Proudfoot, N.J. (1987). Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA 3′ end formation. Cell *49*, 399–406.

Goodwin, E.C., and Rottman, F.M. (1992). The 3′-flanking sequence of the bovine growth hormone gene contains novel elements required for efficient and accurate polyadenylation. J. Biol. Chem. *267*, 16330–16334.

Hargrove, J.L., and Schmidt, F.H. (1989). The role of mRNA and protein stability in gene expression. FASEB J. *3*, 2360–2370.

Harris, H. (1970). The Principles of Human Biochemical Genetics (New York: American Elsevier Publishing Company, Inc.).

Hatada, S., Kuziel, W., Smithies, O., and Maeda, N. (1999). The influence of chromosomal location on the expression of two transgenes in mice. J. Biol. Chem. *274*, 948–955.

Holcik, M., and Liebhaber, S.A. (1997). Four highly stable eukaryotic mRNAs assemble 3′ untranslated region RNA-protein complexes sharing cis and trans components. Proc. Natl. Acad. Sci. USA *94*, 2410–2414.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature *326*, 292–295.

Huang, L.Y., Tholanikunnel, B.G., Vakalopoulou, E., and Malbon,

C.C. (1993). The M(r) 35,000 beta-adrenergic receptor mRNA-binding protein induced by agonists requires both an AUUUA pentamer and U-rich domains for RNA recognition. J. Biol. Chem. *268*, 25769–25775.

Kim, H.S., Lee, G., John, S.W., Maeda, N., and Smithies, O. (2002). Molecular phenotyping for analyzing subtle genetic effects in mice: application to an angiotensinogen gene titration. Proc. Natl. Acad. Sci. USA *99*, 4602–4607.

Koller, B.H., Hagemann, L.J., Doetschman, T., Hagaman, J.R., Huang, S., Williams, P.J., First, N.L., Maeda, N., and Smithies, O. (1989). Germ-line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells. Proc. Natl. Acad. Sci. USA *86*, 8927–8931.

Krege, J.H., Kim, H.S., Moyer, J.S., Jennette, J.C., Peng, L., Hiller, S.K., and Smithies, O. (1997). Angiotensin-converting enzyme gene mutations, blood pressures, and cardiovascular homeostasis. Hypertension *29*, 150–157.

Kubota, N., Terauchi, Y., Miki, H., Tamemoto, H., Yamauchi, T., Komeda, K., Satoh, S., Nakano, R., Ishii, C., Sugiyama, T., et al. (1999). PPARγ mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. Mol. Cell *4*, 597–609.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Oliver, P.M., John, S.W., Purdy, K.E., Kim, R., Maeda, N., Goy, M.F., and Smithies, O. (1998). Natriuretic peptide receptor 1 expression influences blood pressures of mice in a dose-dependent manner. Proc. Natl. Acad. Sci. USA *95*, 2547–2551.

Ross, J. (1995). mRNA stability in mammalian cells. Microbiol. Rev. *59*, 423–450.

Sachs, A.B., and Davis, R.W. (1989). The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell *58*, 857–867.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell *46*, 659–667.

Smithies, O., and Kim, H.S. (1994). Targeted gene duplication and disruption for analyzing quantitative genetic traits in mice. Proc. Natl. Acad. Sci. USA *91*, 3612–3615.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Wang, Z., and Kiledjian, M. (2001). Functional link between the mammalian exosome and mRNA decapping. Cell *107*, 751–762.

Zhang, Y., Buchholz, F., Muyrers, J.P., and Stewart, A.F. (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. Nat. Genet. *20*, 123–128.